The role of environmental variables in structuring landscape-scale species distributions in seafloor habitats

Kraan, Casper; Aarts, Geert; van der Meer, Jaap; Piersma, Theunis

Published in:
Ecology

DOI:
10.1890/09-2040.1

Ecology, 91(6), 2010, pp. 1583–1590

The role of environmental variables in structuring landscape-scale species distributions in seafloor habitats

Casper Kraan,[1,3] Geert Aarts,[1] Jaap van der Meer,[1] and Theunis Piersma[1,2]

[1]*Department of Marine Ecology, Royal Netherlands Institute for Sea Research (NIOZ), P.O. Box 59, 1790 AB Den Burg, Texel, The Netherlands*
[2]*Animal Ecology Group, Centre for Ecological and Evolutionary Studies (CEES), University of Groningen, P.O. Box 14, 9750 AA Haren, The Netherlands*

Abstract. Ongoing statistical sophistication allows a shift from describing species' spatial distributions toward statistically disentangling the possible roles of environmental variables in shaping species distributions. Based on a landscape-scale benthic survey in the Dutch Wadden Sea, we show the merits of spatially explicit generalized estimating equations (GEE). The intertidal macrozoobenthic species, *Macoma balthica*, *Cerastoderma edule*, *Marenzelleria viridis*, *Scoloplos armiger*, *Corophium volutator*, and *Urothoe poseidonis* served as test cases, with median grain-size and inundation time as typical environmental explanatory variables. GEEs outperformed spatially naïve generalized linear models (GLMs), and removed much residual spatial structure, indicating the importance of median grain-size and inundation time in shaping landscape-scale species distributions in the intertidal. GEE regression coefficients were smaller than those attained with GLM, and GEE standard errors were larger. The best fitting GEE for each species was used to predict species' density in relation to median grain-size and inundation time. Although no drastic changes were noted compared to previous work that described habitat suitability for benthic fauna in the Wadden Sea, our predictions provided more detailed and unbiased estimates of the determinants of species–environment relationships. We conclude that spatial GEEs offer the necessary methodological advances to further steps toward linking pattern to process.

Key words: cluster models; GEE; intertidal macrozoobenthos; inundation time; landscape-scale; monitoring; Moran's I; sediment; spatial autocorrelation; Wadden Sea.

Introduction

Spatial autocorrelation: trouble or new paradigm? (Legendre 1993) was an eye-opener in the early 1990s. It marked the beginning of a new framework that underlined the key role of spatial heterogeneity in both statistical analyses and ecological theory (e.g., Wagner and Fortin 2005). Many researchers now routinely include spatial patterning of their focal species as an important ecological variable (e.g., Aarts et al. 2008, Kraan et al. 2009*a*, *b*). Now that spatial autocorrelation is accepted as a general phenomenon (Sparrow 1999, Wagner and Fortin 2005), interest has shifted from pattern to process, i.e., the role of environmental variables in shaping species distributions (Keitt et al. 2002, Dormann et al. 2007).

In marine ecosystems, a range of methods such as canonical correlation (Van der Meer 1999) and logistic-regression (Ysebaert et al. 2002, Compton et al. 2009) have been used to describe macrozoobenthos–environment relationships. However, none of these methods encompass spatial structure in the distribution of either explanatory or response variables. Autocorrelation violates the assumption of independent errors, leading to underestimation of parameter standard errors and biased parameter estimates and model fits (Tognelli and Kelt 2004, Dormann 2007; but see Diniz-Filho et al. 2007). It may lead to flawed conclusions (Lennon 2000, Keitt et al. 2002); Kühn (2007) even observed an inversion of regression slopes between plant species richness and altitude. Therefore, in the likely occurrence of autocorrelation when dealing with spatial data sets, spatially explicit methods are a necessity.

Based on a large-scale benthic research program in the western Dutch Wadden Sea (Kraan et al. 2009*a*, *b*, Van Gils et al. 2009) we highlight the landscape-scale distributions of *Macoma balthica*, *Cerastoderma edule*, *Marenzelleria viridis*, *Scoloplos armiger*, *Corophium volutator*, and *Urothoe poseidonis* (Appendix A) and the environmental variables that structure their spatial patterns. It is generally assumed that large-scale spatial patterns are a response to environmental variables, whereas small-scale spatial patterns are related to biological interactions (Thrush 1991, Legendre 1993). Since we study animal–environment relationships at a landscape-scale, we expect to find a strong decrease of spatial structuring in the distribution of our benthic species after accounting for environmental variability.

Manuscript received 1 November 2009; accepted 5 January 2010. Corresponding Editor: M. Fortin.
[3] E-mail: Casper.Kraan@nioz.nl

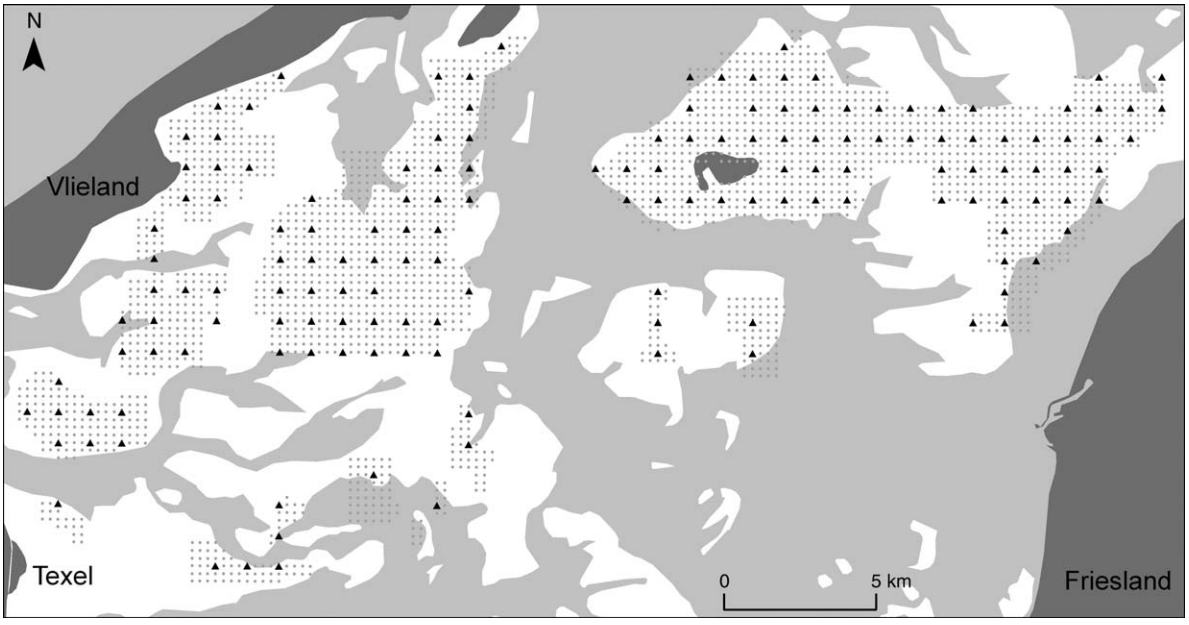

FIG. 1. Map of all benthic (circles) and sediment (triangles) sampling stations, on a 250-m grid and 1000-m grid, respectively, in the western Dutch Wadden Sea. White areas indicate mudflats exposed during low water, intermediate gray areas indicate water, and land is represented by the darkest gray areas.

To handle autocorrelation in response and explanatory variables, as well as non-Gaussian distributions, we apply "generalized estimating equations" (GEE; Liang and Zeger 1986). These are best described as models extending generalized linear models (GLM) with a spatial variance–covariance matrix. We illustrate the impact of autocorrelation on the results for spatially naive GLMs and GEEs with varying assumptions about the correlation structure. In addition to these methodological puzzles, we demonstrate how spatially explicit methods can be used to advance our understanding of species distributions.

METHODS

Study area.—The western Dutch Wadden Sea (53° N, 4–5° E) mainly consists of intertidal- and subtidal mudflats, and gullies, and is bordered to the west and north by the barrier islands of Texel, Vlieland and Terschelling and to the south by the mainland coast of Friesland. About three-quarters of this area consist of soft sediment flats exposed at low tides. Our study area of about 225 km^2 is covered with a grid of fixed sampling stations at 250-m intervals (Fig. 1).

Benthic sampling.—Intertidal macrozoobenthos was sampled in July–September 2005 by rubber boat during high tide or on foot during low tide. Sampling locations, 2784 in total, were found with handheld GPS (Garmin 45 and 12, using WGS84 as map datum [Garmin, Olathe, Kansas, USA]) and at each station 1/56 m^2 was sampled to a depth of 20–25 cm. The sampling cores were sieved over a 1-mm mesh and all individuals were counted and identified (Van Gils et al. 2006*a*, *b*, Kraan et al. 2009*a*, *b*).

Environmental variables.—Sediment samples in a fixed 1000 m grid were also taken (Fig. 1), consisting of sediment from the top 5 cm of the seabed. These samples, 157 in total, were used to determine sediment characteristics, such as median grain-size (MGS; see Bocher et al. 2007). To assign a sediment value to each benthic sampling station, MGS values were interpolated across the whole study area by means of inverse distance weighting (Compton et al. 2009).

Elevation (cm) relative to Dutch Ordinance Level (see Van Gils et al. 2006*b*) was appointed to the nearest sampling station. Furthermore, by comparing elevation with values obtained from an ideal tide with a symmetrical sinusoidal curve, inundation time (IT) per tidal cycle of 12.25 h was calculated. We based calculations on an ideal tide from Harlingen situated on the mainland coast of Friesland (Fig. 1). The number of explanatory variables was limited to MGS, IT, their quadratic terms, as well as their interaction. Other environmental variables, such as distance to gully, were highly collinear with the before mentioned variables and therefore excluded to obtain the most parsimonious model.

Response variables.—Response variables were species-counts (n per sample). From three groups of benthic fauna we have chosen the two most abundant species (Appendix A). (1) Bivalves consisted of *M. balthica* and *C. edule*. *M. balthica* was divided in adults ($\geq$1 growth ring) and juveniles (no growth rings), since different habitat preferences are expected because of juvenile

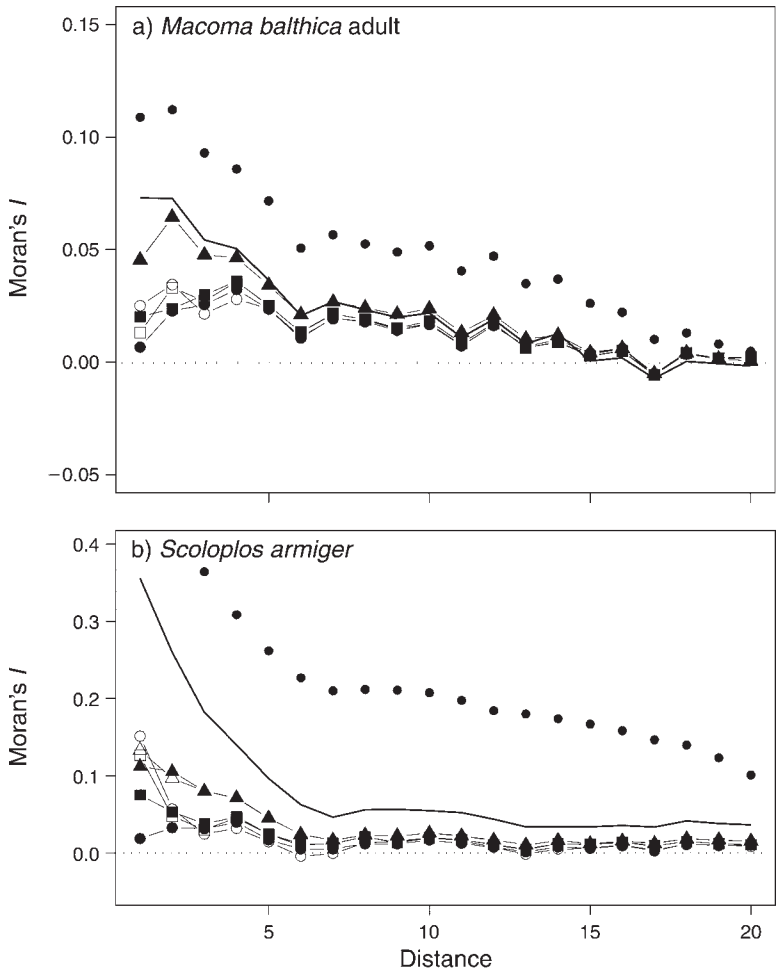

FIG. 2. Residual autocorrelation (Moran's I) of differently structured generalized estimating equations (GEEs) of the full model illustrated for (a) *Macoma balthica* adult and (b) *Scoloplos armiger*. The symbols represent: intercept model (separate dots), GLM or independent correlation structure (solid line), quadratic 4×4 (solid circle), quadratic 3×3 (solid square), quadratic 2×2 (solid triangle), exchangeable 4×4 (open circles), exchangeable 3×3 (open squares), and exchangeable 2×2 (open triangles). Distance intervals on the x-axis are per 250-m interval ($1 = 250$ m, $2 = 500$ m, and so on).

migration (e.g., Beukema 1993). (2) Polychaetes entailed *M. viridis* and *S. armiger*. (3) Crustaceans comprised *C. volutator* and *U. poseidonis*.

Spatial modeling.—Because of the presence of spatial autocorrelation, as shown previously for intertidal benthic fauna in the Dutch Wadden Sea (Kraan et al. 2009*a*, *b*), we opted for GEEs to account for autocorrelation in both the spatial distribution of benthic species and the explanatory environmental variables. Fitting GEEs involves a series of steps. (1) First a GLM assuming a Poisson error distribution was fitted under the assumption that species counts were independent from each other. (2) This GLM was used to estimate the spatial correlation matrix of the residuals. (3) The regression model is then extended by incorporating the spatial correlation matrix. Prior to these iterative steps (2 and 3) the data were clustered into smaller blocks of 2×2, 3×3, or 4×4 sampling stations to reduce computation time. Correlations within each cluster were included in the model, while correlations between clusters were assumed to be absent (Carl and Kühn 2007, Koper and Manseau 2009).

The correlation within a cluster can be specified by two different structures: (1) *quadratic*, where the correlation varies with distance class (Carl and Kühn 2007; for example, a block of 2×2 sampling stations contains two distance classes, therefore two correlation parameters need to be estimated) and (2) *exchangeable*. All correlations within a cluster are equal. A third structure, *fixed*, where the correlation decreases with distance and can be estimated by Moran's I of GLM residuals (therefore clustering is not needed [Dormann et al. 2007]), was omitted. This structure may lead to contradicting results, such as a large reduction of residual autocorrelation combined with, on occasion, smaller estimated standard errors than the nonspatial

TABLE 1. Selection of the best correlation structure (bold), i.e., independence (ind.), quadratic (quad.), or exchangeable (exch.), and best cluster size illustrated for *Macoma balthica* adult and *Scoloplos armiger* based on the residual autocorrelation (Moran's *I*) at distance class 1 (250 m).

Model type	QIC	Quasi-likelihood	Trace	Moran's *I* at class 1	Moran's *I* at class 5
Macoma balthica					
GEE ind.	2558.94	−1275.01	4.46	0.09	0.06
GEE quad. 4 × 4	**2567.97**	**−1276.25**	**7.74**	**0.01**	**0.02**
GEE quad. 3 × 3	2565.41	−1275.86	6.84	0.02	0.03
GEE quad. 2 × 2	2560.53	−1275.06	5.20	0.05	0.05
GEE exch. 4 × 4	2568.72	−1276.37	7.99	0.03	0.03
GEE exch. 3 × 3	2565.33	−1275.84	6.83	0.01	0.03
GEE exch. 2 × 2	2560.52	−1275.06	5.20	0.05	0.05
Scoloplos armiger					
GEE ind.	−766.10	386.73	3.68	0.31	0.17
GEE quad. 4 × 4	**−660.73**	**339.95**	**9.58**	**0.02**	**0.03**
GEE quad. 3 × 3	−686.11	350.42	7.36	0.07	0.05
GEE quad. 2 × 2	−742.41	376.95	5.74	0.11	0.08
GEE exch. 4 × 4	−722.99	371.42	9.93	0.15	0.06
GEE exch. 3 × 3	−688.41	351.67	7.47	0.13	0.05
GEE exch. 2 × 2	−741.19	376.35	5.75	0.13	0.09

Note: Key to abbreviations: GEE, generalized estimating equations; QIC, quasi-likelihood under the independence model information criterion.

models. Another downside is the long computation time and large matrices, often failing convergence (Dormann et al. 2007). More details of GEEs and comparison with other methods see Hardin and Hilbe (2003) and Dormann et al. (2007).

To choose the most appropriate correlation structure and cluster size, we visually compared the correlograms (e.g., Kraan et al. 2009*a*) of the residuals of the full GLM and Pearson residuals of the GEEs standardized by the working correlation (e.g., Carl and Kühn 2007; Fig. 2). We selected the correlation structure and cluster size leading to the lowest small-scale (i.e., 250-m) residual autocorrelation (e.g., Carl and Kühn 2007, Dormann et al. 2007). Next task was to obtain the minimal adequate model. GEE uses a quasi-likelihood framework for model estimation (Liang and Zeger 1986); therefore, a maximum-likelihood based method for model selection such as Akaike's Information Criterion is not valid. Instead, we used backward model selection based on the quasi-likelihood under the independence model information criterion, or QIC (Pan 2001):

$$\text{QIC} = -2Q + 2\text{trace}(\hat{\Omega}_I^{-1}\hat{\mathbf{V}}_R)$$

where Q is the quasi-likelihood calculated as $Q = y \times \log \mu - \mu$ where μ represents the mean (McCullagh and Nelder 1989; Appendix C). The p-dimensional matrices $\hat{\mathbf{\Omega}}_I$ and $\hat{\mathbf{V}}_R$ are variance estimators of the regression coefficients under the correlation structures I (independence) and R (GEE-based), respectively (Cui and Feng 2009).

In the presence of spatial autocorrelation, we expect, under the model (incorrectly) assuming independent data, an underestimation of the parameter variances. Accounting for spatial autocorrelation will most likely lead to higher parameter variances and hence a larger trace ($\hat{\Omega}_I^{-1}\hat{\mathbf{V}}_R$). Like Akaike's information criterion, the model with the lowest QIC is the best model (Pan 2001). Note that QIC is not suitable for selecting the most appropriate correlation structure (Koper and Manseau 2009; but see Hin et al. 2007). QIC, by definition, favors those models with the lowest trace ($\hat{\Omega}_I^{-1}\hat{\mathbf{V}}_R$). However, it does allow determining the best subset of covariates (Cui 2007). In general, model selection under spatial autocorrelation has still to be developed and no general guidelines exist yet (Dormann et al. 2007, Koper and Manseau 2009). All analyses were done using R (version 2.6.2; R Development Core Team 2008; see Supplement for code) following Carl and Kühn (2007).

RESULTS

A clear dichotomy existed between residual patterns of spatial and nonspatial models describing species–environment relationships in intertidal areas of the western Dutch Wadden Sea. Intercept models (species counts only) and independent correlation structures, which approximate spatially naïve GLMs, contained residual patterning (Fig. 2, Table 1). Spatial models reduced spatial autocorrelation (Fig. 2, Table 1).

Overall, an independent correlation structure was under no circumstances selected as the best correlation structure to model the spatial interaction between species and explanatory environmental variables (Table 1). Also, a 2 × 2 cluster size never appeared to be the best cluster size (Table 1). The 4 × 4 quadratic correlation structure was selected for all species, except for *C. volutator* (3 × 3 exchangeable) and *U. poseidonis* (3 × 3 quadratic). For *C. edule*, we illustrated the effects of accounting for spatial autocorrelation (Appendix C), by comparing a spatially naïve GLM and a spatially explicit GEE, assuming a quadratic correlation structure and a 4 × 4 cluster size. Indeed, as expected, the standard errors became much larger (Appendix C). Consequently, the absolute values of all GEE parame-

ters, except the intercept, became smaller. This showed that ignoring spatial autocorrelation not only leads to a different assessment of the significance of an environmental variable, it also illustrated that slopes became flatter.

Given the best correlation structure and cluster size (Table 1), we used backward model selection to determine the importance of the environmental variables IT and MGS, their quadratic terms and interactions (Table 2). If deleting a variable lowered the QIC, this suggested a better model than the full model with which we started. In practice, however, only the interaction MGS × IT for the *U. poseidonis* model was excluded (Table 2).

The best model for each species was used to predict species' density or "perceived preference" (preference subsequently) for the environmental variables MGS and IT. Predictions were only made for IT and MGS values larger than the 0.01 quantile and smaller than the 0.99 quantile (Appendix B). These restrictions downsized potential over-emphasis of scarce and unlikely MGS and IT combinations. The analysis showed highest densities for both adult and juvenile *M. balthica* in muddy sandflats with a short to medium IT (Fig. 3a, b). In addition, adults also preferred coarse sediments with a long IT (Fig. 3a). *C. edule* preferred a wide range of MGS with an IT of 6–8 h (Fig. 3c). The spionid *M. viridis* preferred a particularly narrow range of environmental characteristics, i.e., a MGS range of 100–150 μm and an IT of 8–10 h (Fig. 3d). *S. armiger* mainly preferred coarse sediments over the complete range of available IT (Fig. 3e). *C. volutator* (Fig. 3f) preferred muddy sediments with an IT of 6–8 h. *U. poseidonis* combined a preference for coarse sediments with intermediate IT (Fig. 3g).

TABLE 2. Model selection, to obtain the minimal adequate model (bold), based on QIC values.

Model	QIC
Macoma balthica adult	
Full	**2567.97**
−IT	2572.91
−MGS^2	2587.68
−MGS	2615.96
−MGS × IT	2617.13
−IT^2	2677.24
Macoma balthica juvenile	
Full	**1539.82**
−IT	1540.78
−MGS^2	1543.26
−MGS × IT	1549.65
−IT^2	1840.45
−MGS	2189.62
Cerastoderma edule	
Full	**3066.09**
−MGS^2	3068.26
−MGS	3110.07
−IT	3137.24
−IT^2	3165.55
−MGS × IT	4216.42
Marenzelleria viridis	
Full	**−15 420.0**
−MGS × IT	−15 385.9
−IT	−14 206.6
−IT^2	−14 011.9
−MGS	−10 883.1
−MGS^2	598.2
Scoloplos armiger	
Full	**−660.73**
−MGS × IT	−649.65
−IT	−428.52
−IT^2	−132.43
−MGS^2	587.49
−MGS	2092.86
Corophium volutator	
Full	**−44 938.5**
−MGS × IT	−44 800.4
−IT	−44 286.5
−MGS	−43 200.7
−IT^2	−35 830.1
−MGS^2	−7495.29
Urothoe poseidonis	
Full	−17 719.7
−MGS × IT	**−17 776.3**
−IT	−15 731.9
−IT^2	−14 088.5
−MGS^2	−11 617.6
−MGS	−6926.49

Notes: Shown is model improvement after backward selection of an explanatory variable or an interaction between variables. Variables increase in importance from top to bottom. For each species the best correlation structure and cluster size were applied.

DISCUSSION

In marine ecosystems, thus far, species–environment relationships have been analyzed with methods that do not take spatial autocorrelation into account. Although this does not necessarily imply that these analyses are flawed (Diniz-Filho et al. 2003; but see Lennon 2000), some bias in model coefficients (Dormann 2007, Bini et al. 2009) and decrease in model parsimony probably did occur. Our results support that species–environment relationships should be done with spatially explicit methods, such as GEE or alternatives such as generalized linear mixed models (GLMM, Pinheiro and Bates 2000) or Bayesian model-based geostatistics (Diggle and Ribeiro 2007). The kind of correction appears to be of minor importance, as long as a correction is made (see *Results*; Keitt et al. 2002). However, efficiency gains can still be made by using the best correlation structure and best cluster size, followed by model "fine-tuning" based on QIC (Fig. 2).

Thus far, the number of studies applying spatial GEEs is very limited (Carl and Kühn 2007); our study constitutes only the second application on a landscape scale with survey data and a large number of samples, and it is the first in a marine setting. GLMMs or model-based geostatistics would be equally suitable for these kinds of spatial analyses (Dormann et al. 2007); however, the size of the data set is a limiting step, which is circumvented by GEEs that slice the data into clusters. Note that these latter methods are better suited

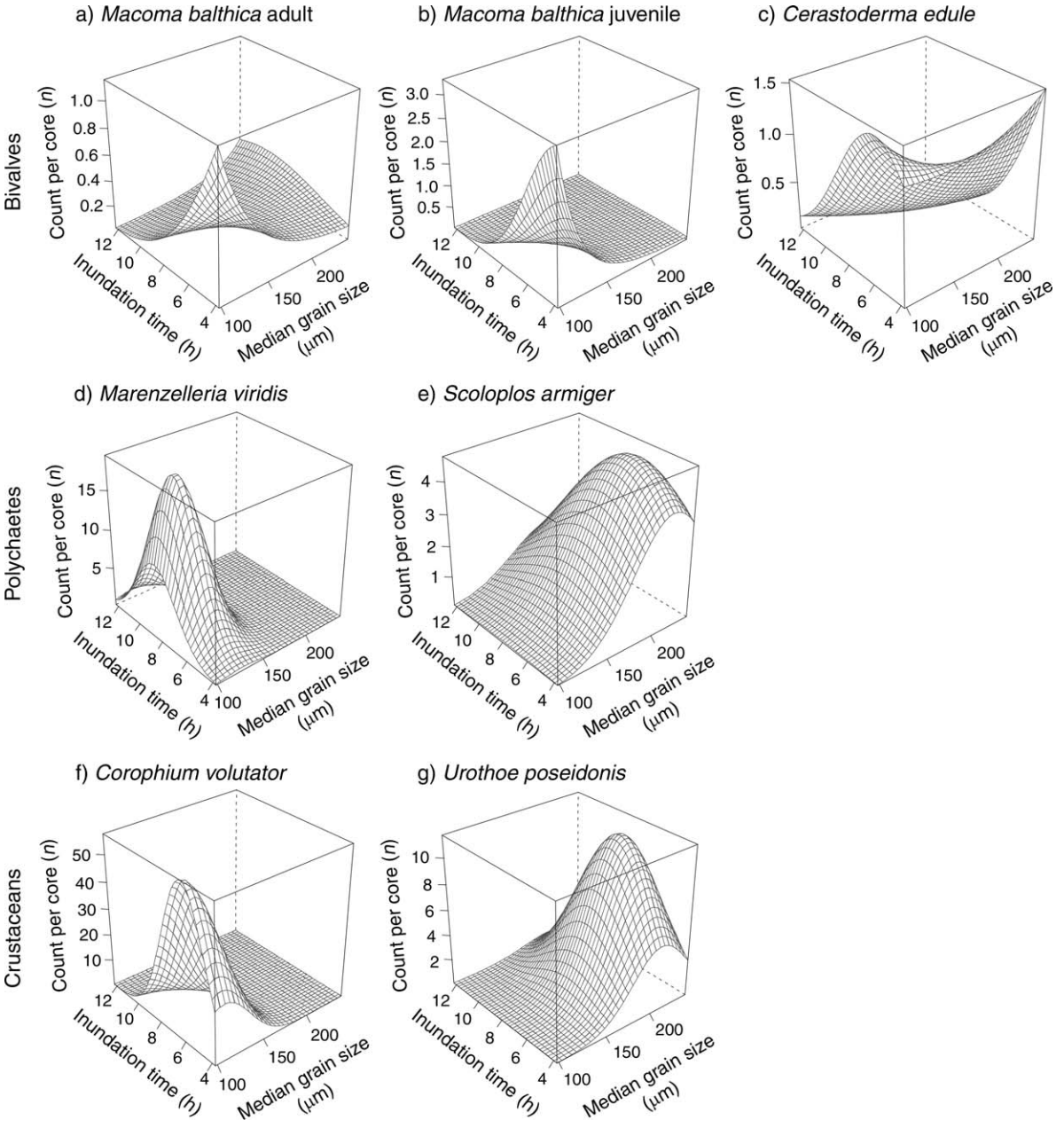

FIG. 3. Fitted preference of macrobenthic species in relation to median grain size (MGS) and inundation time (IT) in the western Dutch Wadden Sea. For each species the minimal adequate GEE was applied (Table 2).

for spatial prediction and interpolation than GEEs because spatially correlated errors can be predicted for non-sampled stations (Diggle and Ribeiro 2007, Dormann et al. 2007).

As hypothesized, residual structure was greatly reduced after accounting for environmental variability. Indeed, the landscape-scale distribution of hidden marine intertidal species in the western Dutch Wadden Sea could, for a larger part, be attributed to IT and MGS (Tables 1 and 2). The remaining residual spatial variation, which represents variation not accounted for by the selected physical parameters, can be credited to either an environmental variable not included in the analyses or biotic interactions between or within species (Keitt et al. 2002, Wagner and Fortin 2005). Disentangling these is an important goal (Wagner and Fortin 2005, Dormann et al. 2007). In this respect, dynamic regressions that estimate competition coefficients could be rewarding (Pfister 1995), especially if this approach would be extended with a variance–covariance matrix to include autocorrelation. In addition, there remains a possibility that the included environmental variables are

driven by other, unknown, spatially patterned factors (Keitt et al. 2002, Diniz-Filho et al. 2003). However, this feature is shared by all regression-type analyses and can not be solved without experiments. Nevertheless, acknowledging and accounting for spatial autocorrelation, as presented here, is a marked improvement from spatially naïve analyses, because better models can be derived and pinpoints that ecology should encompass spatial autocorrelation (e.g., Sparrow 1999, Wagner and Fortin 2005).

The spatially explicit framework applied for Wadden Sea benthic fauna likely offers precise and statistically sound estimates of environmental variables governing species abundances, compared to spatially naïve approaches (Appendix C). The question remains: actually how robust are previous spatially naïve analyses of species–environment associations in this area? Unfortunately, no such analyses at a landscape scale overlap the extent of our study area. Therefore, we can not evaluate former results and determine if drastic differences occur, as previously shown by Tognelli and Kelt (2004) and Kühn (2007). However, habitat associations of macrobenthic fauna in marine intertidal areas have been particularly centered on the use of sediment grain-size distributions to approximate habitat suitability (e.g., Ysebaert et al. 2002, Compton et al. 2009). Indeed, the present study also indicates this is the most important habitat characteristic (Table 2). Yet, IT also contributed significantly to habitat suitability for benthic fauna, as did the interaction between IT and MGS (Fig. 3, Table 2). Without applying spatially explicit analyses, such ecological patterns would leave room for discussion about the correct relationships between variables.

Ignoring the different scales of research, a few studies offer insight in species–environment relationships in the Wadden Sea, and allow qualitative comparisons of habitat preferences. Some examples: (1) similar to our results (Fig. 3b), others (e.g., Beukema 1993) show that juvenile *M. balthica* prefer short IT, areas high in the intertidal zone. (2) Because of their peculiar downslope winter migration (Beukema 1993), adult *M. balthica* should occur in the middle and lower zones in the intertidal (e.g., Beukema 1993); this is only partially apparent in our results (Fig. 3a). Our more recent analyses seem to suggest that a large part of the adults do not migrate from the shallow and muddy areas toward the deeper and sandier regions of the Wadden Sea. Given the rapid decline of the population in the western Dutch Wadden Sea (Van Gils et al. 2009), their habitat preferences might have shifted in recent times.

To sum up, our results do not show drastic changes, but indicate that we achieved a more detailed and unbiased insights in the determinants of species–environment relationships in the intertidal ecosystem of the Dutch Wadden Sea. This spatial approach extents current knowledge, and offers enhanced understanding of species distributions, which still is underexplored in marine science (Compton et al. 2009, Thrush et al. 2009). GEEs offered the necessary methodological advances to describe species distributions in a spatially patterned environment.

Acknowledgments

During most of the fieldwork, led by A. Dekinga, we were based on the Royal NIOZ-research vessel RV *Navicula*, and we thank K. van de Star, T. van der Vis, and H. de Vries for their help. We thank Vereniging Natuurmonumenten for permission to work around the island of Griend. A large number of volunteers and students contributed to the collection of the field data. We also thank T. Compton for help with calculating inundation time.

Literature Cited

Aarts, G., M. MacKenzie, B. McConnell, M. Fedak, and J. Matthiopoulos. 2008. Estimating space-use and habitat preference from wildlife telemetry data. Ecography 31:140–160.

Beukema, J. J. 1993. Successive changes in distribution patterns as an adaptive strategy in the bivalve *Macoma balthica* (L.) in the Wadden Sea. Helgoländer Meeresuntersuchungen 47: 287–304.

Bini, L. M., et al. 2009. Coefficient shifts in geographical ecology: an empirical evaluation of spatial and non-spatial regression. Ecography 32:193–204.

Bocher, P., T. Piersma, A. Dekinga, C. Kraan, M. G. Yates, T. Guyot, E. O. Folmer, and G. Radenac. 2007. Site- and species-specific distribution patterns of molluscs at five intertidal soft-sediment areas in northwest Europe during a single winter. Marine Biology 151:577–594.

Carl, G., and I. Kühn. 2007. Analyzing spatial autocorrelation in species distributions using Gaussian and logit models. Ecological Modelling 207:159–170.

Compton, T. J., T. A. Troost, J. Drent, C. Kraan, P. Bocher, J. Leyrer, A. Dekinga, and T. Piersma. 2009. Repeatable sediment associations of burrowing bivalves across six European tidal flat systems. Marine Ecology Progress Series 382:87–98.

Cui, J. 2007. QIC program and model selection in GEE analyses. Stata Journal 7:209–220.

Cui, J., and L. Feng. 2009. Correlation structure and model selection for negative binomial distribution in GEE. Communications in Statistics: Simulation and Computation 38: 190–197.

Diggle, P. J., and P. J. Ribeiro, Jr. 2007. Model-based geostatistics. Springer, New York, New York, USA.

Diniz-Filho, J. A. F., L. M. Bini, and B. A. Hawkins. 2003. Spatial autocorrelation and red herrings in geographical ecology. Global Ecology and Biogeography 12:53–64.

Diniz-Filho, J. A. F., B. A. Hawkins, L. M. Bini, P. De Marco, Jr., and T. Blackburn. 2007. Are spatial regression methods a panacea or a Pandora's box? A reply to Beale et al. (2007). Ecography 30:848–851.

Dormann, C. 2007. Effects of incorporating spatial autocorrelation into the analysis of species distribution data. Global Ecology and Biogeography 16:129–138.

Dormann, C., et al. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography 30:609–628.

Hardin, J. W., and J. M. Hilbe. 2003. Generalized estimating equations. Chapman and Hall/CRC, Boca Raton, Florida, USA.

Hin, L.-Y., V. J. Carey, and Y.-G. Wang. 2007. Criteria for working-correlation-structure selection in GEE: assessment via simulation. American Statistician 61:360–364.

Keitt, T. H., O. N. Bjørnstad, P. M. Dixon, and S. Citron-Pousty. 2002. Accounting for spatial pattern when modeling organism–environment interactions. Ecography 25:616–625.

Koper, N., and M. Manseau. 2009. Generalized estimating equations and generalized linear mixed-effects models for modelling resource selection. Journal of Applied Ecology 46: 590–599.

Kraan, C., J. van der Meer, A. Dekinga, and T. Piersma. 2009*a*. Patchiness of macrobenthic invertebrates in homogenized intertidal habitats: hidden spatial structure at a landscape scale. Marine Ecology Progress Series 383:211–224.

Kraan, C., J. A. van Gils, B. Spaans, A. Dekinga, A. I. Bijleveld, M. van Roomen, R. Kleefstra, and T. Piersma. 2009*b*. Landscape-scale experiment demonstrates that Wadden Sea intertidal flats are used to capacity by molluscivore migrant shorebirds. Journal of Animal Ecology 78:1259–1268.

Kühn, I. 2007. Incorporating spatial autocorrelation may invert observed patterns. Diversity and Distributions 13:66–69.

Legendre, P. 1993. Spatial autocorrelation: trouble or new paradigm? Ecology 74:1659–1673.

Lennon, J. J. 2000. Red-shifts and red herrings in geographical ecology. Ecography 23:101–113.

Liang, K.-Y., and S. L. Zeger. 1986. Longitudinal data analysis using generalized linear models. Biometrika 73:13–22.

McCullagh, P., and J. A. Nelder. 1989. Generalized linear models. Chapman and Hall, London, UK.

Pan, W. 2001. Akaike's information criterion in generalized estimating equations. Biometrics 57:120–125.

Pfister, C. A. 1995. Estimating competition coefficients from census data: a test with field manipulations of tidepool fishes. American Naturalist 146:271–291.

Pinheiro, J. C., and D. M. Bates. 2000. Mixed-effects models in S and S-plus. Springer-Verlag, New York, New York, USA.

R Development Core Team. 2008. R: a language and environment for statistical computing. R Project for Statistical Computing, Vienna, Austria. ⟨http://www.R-project.org⟩

Sparrow, A. D. 1999. A heterogeneity of heterogeneities. Trends in Ecology and Evolution 14:422–423.

Thrush, S. F. 1991. Spatial patterns in soft-bottom communities. Trends in Ecology and Evolution 6:75–79.

Thrush, S. F., J. E. Hewitt, P. K. Dayton, G. Coco, A. M. Lohrer, A. Norkko, J. Norkko, and M. Chiantore. 2009. Forecasting the limits of resilience: integrating empirical research with theory. Proceedings of the Royal Society B 276: 3209–3217.

Tognelli, M. F., and D. A. Kelt. 2004. Analysis of determinants of mammalian species richness in South America using spatial autoregressive models. Ecography 27:427–436.

Van der Meer, J. 1999. Keeping things in order: multivariate direct gradient analysis of a strongly fluctuating benthic community. Journal of Sea Research 42:263–273.

Van Gils, J. A., C. Kraan, A. Dekinga, A. Koolhaas, J. Drent, P. de Goeij, and T. Piersma. 2009. Reversed optimality and predictive ecology: burying depth forecasts population change in a bivalve. Biology Letters 5:5–8.

Van Gils, J. A., T. Piersma, A. Dekinga, B. Spaans, and C. Kraan. 2006*a*. Shellfish-dredging pushes a flexible avian top predator out of a protected marine ecosystem. PLoS Biology 4:2399–2404.

Van Gils, J. A., B. Spaans, A. Dekinga, and T. Piersma. 2006*b*. Foraging in a tidally structured environment by red knots (*Calidris canutus*): ideal, but not free. Ecology 87:1189–1202.

Wagner, H. H., and M.-J. Fortin. 2005. Spatial analyses of landscapes: concepts and statistics. Ecology 86:1975–1987.

Ysebaert, T., P. Meire, P. M. J. Herman, and H. Verbeek. 2002. Macrobenthic species response surfaces along estuarine gradients: prediction by logistic regression. Marine Ecology Progress Series 225:79–95.

APPENDIX A

Minimum and maximum (mean) counts per sampling station of the benthic species (*Ecological Archives* E091-110-A1).

APPENDIX B

Scatterplot of median grain size (MGS) and inundation time (IT) (*Ecological Archives* E091-110-A2).

APPENDIX C

Comparison between regression coefficients and robust standard errors of a quasi-Poisson generalized linear model (GLM) and a generalized estimating equation (GEE) with a quadratic 4×4 correlation structure for the minimal adequate model for *Cerastoderma edule* (*Ecological Archives* E091-110-A3).

SUPPLEMENT

Script for running GEEs in R and using QIC values to select the best model (*Ecological Archives* E091-110-S1).